This paper is available on line at http://www.mcponline.org

Quantitative Assessment of Chromatin Immunoprecipitation Grade Antibodies Directed against Histone Modifications Reveals Patterns of Co-occurring Marks on Histone Protein Molecules*[S]

Sally E. Peach, Emily L. Rudomin, Namrata D. Udeshi, Steven A. Carr, and Jacob D. Jaffe‡

The defining step in most chromatin immunoprecipitation (ChIP) assays is the use of an antibody to enrich for a particular protein or histone modification state associated with segments of chromatin. The specificity of the antibody is critical to the interpretation of the experiment, yet this property is rarely reported. Here, we present a quantitative method using mass spectrometry to characterize the specificity of key histone H3 modification-targeting antibodies that have previously been used to characterize the "histone code." We further extend the use of these antibody reagents to the observation of long range correlations among disparate histone modifications. Using purified human histones representing the mixture of chromatin states present in living cells, we were able to quantify the degree of target enrichment and the specificity of several commonly used, commercially available ChIP grade antibodies. We found significant differences in enrichment efficiency among various reagents directed against four frequently studied chromatin marks: H3K4me2, H3K4me3, H3K9me3, and H3K27me3. For some antibodies, we also detected significant off target enrichment of alternate modifications at the same site (*i.e.*, enrichment of H3K4me2 by an antibody directed against H3K4me3). Through cluster analysis, we were able to recognize patterns of co-enrichment of marks at different sites on the same histone protein. Surprisingly, these co-enrichments corresponded well to "canonical" chromatin states that are exemplary of activated and repressed regions of chromatin. Altogether, our findings suggest that 1) the results of ChIP experiments need to be evaluated with caution given the potential for cross-reactivity of the commonly used histone modification recognizing antibodies, 2) multiple marks with consistent biological interpretation exist on the same histone protein molecule, and 3) some components of the histone code may be transduced on single proteins in living cells. ***Molecular & Cellular Proteomics 11: 10.1074/mcp.M111.015941, 128–137, 2012.***

The core histone proteins are known to harbor dozens of post-translational modifications including methylation, acetylation, and phosphorylation. It has been hypothesized that these modifications regulate access to chromatin in a locus-specific manner allowing for selective transcription, repression, and silencing of genomic regions (1).

The chromatin immunoprecipitation (ChIP)[1] assay has emerged as a technique for genome-wide studies of histones and the distributions of their modifications (2, 3). Such studies have contributed to our understanding of the functional role of specific post-translational modifications or "marks." For example, it was recognized that H3K4me3 marks often localize near the promoters of transcriptionally active genes, whereas H3K9me3 localizes to inactive heterochromatic regions of the genome (4). More recently, so called "bivalent domains" were recognized in ES cells: genomic loci that bear marks of activation (H3K4me3) and repression (H3K27me3) simultaneously that seem to resolve to one of the two states, either solely active or repressed, during differentiation (5). It is important to point out that the simultaneous presence of these marks can be localized to a region, but not necessarily the same exact histone molecule.

The defining step in most ChIP assays is the use of an antibody to enrich for a particular protein or modification state associated with segments of chromatin. The specificity of the antibody is critical to the interpretation of the experiment, yet this property is rarely, if ever, reported. However, the interpre-

From the Broad Institute, Proteomics Platform, Cambridge, Massachusetts 02142

Received November 16, 2011, and in revised form, February 28, 2012

Published, MCP Papers in Press, March 21, 2012, DOI 10.1074/mcp.M111.015941

[1] The abbreviations used are: ChIP, chromatin immunoprecipitation; SILAC, stable isotope labeling of amino acids in cell culture; XIC, extracted ion chromatogram; IP, immunoprecipitation; AIMS, accurate inclusion mass screening.

tation of a ChIP experiment hinges on the faithful enrichment of the target and the exclusion of other moieties. Furthermore, global background frequencies of histone modifications in chromatin govern the limits of signal-to-noise in ChIP experiments and, if known, could aid in improving peak-calling algorithms. However, antibody-based techniques cannot estimate these frequencies.

In principle, antibodies of sufficient selectivity should allow us to answer a key question in chromatin biology: which marks are simultaneously present on the same histone molecule? For example, if an antibody were to purify the population of all H3K4me3-bearing histones to homogeneity, we could definitively answer the question as to whether the bivalent state—H3K4me3/H3K27me3—is ever represented on the same histone protein molecule. However, this is only true if the antibody reagent can perform this purifying role, and assessment of an antibody's performance in this regard has been difficult.

MS is an analytical technique capable of directly probing protein primary structure. It is well suited to examining post-translational modifications such as methylation and acetylation that commonly occur on histone tails (6). Unambiguous assignments of the sites of these common histone marks are easily accomplished by MS. As the name of the technique implies, MS is also an exquisite differentiator of biochemical entities by mass. Stable isotope techniques therefore may be employed to generate test and reference populations of proteins that are nearly chemically indistinguishable in living cells but can be measured independently by a mass spectrometer. Here, we use stable isotope labeling of amino acids in cell culture (SILAC) to quantitatively assess the specificity of various ChIP grade antibodies and extend their use to identify co-occurring marks on histone protein molecules.

EXPERIMENTAL PROCEDURES

HeLa S3 cells were grown in RPMI 1640 SILAC medium (Caisson Laboratories) containing 5% dialyzed FBS (Invitrogen) and either Arg-0 (natural abundances of ^{12}C and ^{14}N), Arg-6 ($^{13}C_6$), or Arg-10 ($^{13}C_6$,$^{15}N_4$). Because the downstream derivatization and digestion strategy results in proteolytic cuts only C-terminal of Arg residues, labeling of Lys was not required. The cells were harvested by centrifugation, washed twice in ice-cold PBS, and treated exactly as in Ref. 7. Briefly, cells were gently lysed, the nuclei were pelleted by centrifugation, and histones were acid extracted using 0.4 N H_2SO_4. After precipitation of histone proteins with 20% trichloroacetic acid, HPLC separation of the histones was performed essentially as in Ref. 7 except that a Zorbax C8 column (4.6 × 150 mm; Agilent Technologies) was employed. The separations were carried out on an Agilent 1100 analytical scale HPLC (Agilent Technologies). All of the peaks corresponding to H3 were pooled, concentrated by vacuum centrifugation and resuspended in ChIP buffer (Millipore 20-153; 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8.1, 167 mM NaCl). A typical yield would be ~50 μg of H3/10^7 cells as measured by the CoomassiePlus protein assay (Pierce).

The antibodies used in this study were: Abcam ab24834 (pan-H3, no particular modification target), ab1012 (anti-H3K4me3), and ab6147 (anti-H3K27me3); Cell Signaling Technologies CST9726 (anti-H3K4me2), CST9751 (anti-H3K4me3), CST9754 (anti-H3K9me3), and CST9733 (anti-H3K27me3); and Millipore MP07-449 (anti-H3K27me3). For each immunoaffinity enrichment experiment, we mixed 25 μl of protein A/G-agarose beads (Pierce 20422) with 2 μg of antibody and 20 μg of H3 in a total volume of 100 μl (any additional volume was made up by ChIP buffer). Incubation proceeded overnight at 4 °C with gentle agitation. After incubation, the beads were washed three times with 100 μl of ChIP buffer and eluted with 0.1% TFA in water. Eluates from three different conditions were pooled at this time; *i.e.*, we mixed the eluates from an enrichment of the pan-specific H3 antibody (Abcam ab24834) using SILAC light histones with eluates from two modification-specific H3 antibodies (*e.g.*, Abcam ab1012 and CST9751) using SILAC medium and heavy histones, respectively. See Fig. 1*A* for schematic details of a simplified two condition comparison (the SILAC "medium" channel is omitted for clarity). We typically pooled 1/10 of the volume of the pan-specific antibody eluate with the eluates from the modification-specific antibodies for normalization purposes, because the yield of the pan-specific IP was generally much higher. The pooled eluate was adjusted to contain 1× NuPAGE loading buffer and separated on a 4–12% NuPAGE gradient gel (Invitrogen).

A band corresponding to the molecular weight of H3 was excised from the gel (although it was not always visible). The gel band was destained and dehydrated according to standard protocols (8) and rehydrated in 100 mM ammonium bicarbonate. We derivatized the protein in the gel by covering the band with a solution of 75% propionic anhydride (Sigma), 25% methanol for 30 min with shaking at room temperature. This derivatization blocks any lysines that would normally be cleaved by trypsin (the unmodified and monomethyl states; dimethyl-lysine, trimethyl-lysine, and acetyl-lysines are not reactive with propionic anhydride nor cleavable by trypsin). The derivatization solution was extensively washed out with several changes of 100 mM ammonium bicarbonate, and then in-gel trypsin digestion proceeded according to standard protocols. After digestion and peptide extraction, the samples were vacuum-centrifuged to dryness. The peptides were derivatized with propionic anhydride again according to Ref. 9. After the second propionylation step, the peptides were vacuum-centrifuged to dryness and resuspended in 5 μl of 3% acetonitrile, 5% formic acid.

The samples were analyzed using an LC-MS system consisting of an Agilent 1100 nanoflow HPLC coupled to an Orbitrap XL mass spectrometer (ThermoFisher Scientific). Chromatography and gradient conditions were the same as in Ref. 10. The peptides were separated on a home-made 13-cm × 75-μm column packed with ReproSil C18 3-μm particles (Dr. Maisch GmbH). The mass spectrometry acquisition method consisted of two segments over the course of a 90-min analysis. The first 20 min were designed to ensure quantification of H3K4me2 and H3K4me3 peptides, which elute early in the experiment. During this segment, we acquired, in the following order: 1) a selected ion monitoring scan from *m/z* 385 to 415, 2) a data-independent MS/MS scan on precursor 394.7, with an isolation width 2.5 *m/z* and normalized CE 30, scanning from *m/z* 105 to 815 (designed to catch H3K4me2 in the "light" SILAC state), 3) a data-independent MS/MS scan on precursor 401.7, with isolation with 2.5 *m/z* and normalized CE 30, scanning from *m/z* 110 to 815 (designed to catch H3K4me3 in the light SILAC state), and 4) a full MS scan from *m/z* 290 to 1700. During the second segment (minutes 20–90), we performed an AIMS experiment (10) where we placed a list of selected precursors on an inclusion list (see supplemental Table 1 for AIMS list). The *m/z* targets on this list were directed at the SILAC light states of the modified peptides. We acquired a selected ion monitoring MS from *m/z* 350 to 800 and selected, in intensity order, the top five precursors found in the selected ion monitoring MS scan from the AIMS list for data-dependent MS/MS scans (normalized CE of 30, isolation width of 3.0 *m/z*). Dynamic exclusion was enabled with a

repeat count of 2, repeat duration of 10 s, and exclusion duration of 10 s, with a width of ± 7.5 ppm. Only AIMS list members could trigger MS/MS scans (most intense if no parent masses found not enabled). Lock mass was not enabled. Spray voltage was 2.2 kV.

All of the spectra were interpreted by manual inspection with the assistance of the MS Product module of Spectrum Mill (Agilent Technologies with in-house modifications). Representative reference spectra for all of the SILAC light state modified peptides can be found in the supplemental Fig. 1 package. We used the XCalibur QualBrowser to generate extracted ion chromatogram (XIC) layouts corresponding to the light, medium, and heavy precursors of all of the peptides that we monitored. The peaks were integrated using the ICIS algorithm, and peak areas for each modification state in the light, medium, and heavy SILAC labels were tabulated in each sample. We always verified the identity of a peak by a corresponding MS/MS scan before accepting it for quantification. The ratio of enrichment by a modification-specific antibody (Arg-6 medium and Arg-10 heavy channels) was computed by dividing the peak area with that of the pan-specific H3 antibody (Arg-0 light channel). All of the ratios were normalized by the medium/light or heavy/light ratio of H3(41–49), which does not contain any lysines and therefore cannot be methylated or acetylated. Each antibody enrichment experiment was repeated two or three times. The results from each experiment as well as averages and standard deviations can be found in supplemental Table 4. Assessment of methylesterification artifacts caused by sample preparation is detailed in supplemental Fig. 2 and was found to be <1% using our methods.

To further improve our estimates of the background percentages of histone modifications, we obtained synthetic peptides corresponding to the peptides in our analysis. An equimolar mixture of all of these synthetic peptides was constructed and analyzed via LC-MS on a Q-Exactive mass spectrometer with LC conditions identical to those above. The mass spectrometer data acquisition consisted of a single full scan followed by data-independent MS/MS HCD scans (CEs vary by peptide) targeting all of the synthetic peptides in the mixture. The peaks were identified using Skyline (11), and MS1 peak areas were extracted for all observable precursor charge states. The relative ionization efficiency of each peptide modification state belonging to the same underlying sequence was determined using MS1 peak areas summed from all observable charge states (supplemental Table 2).

Hierarchical clustering of data was performed and visualized with GenePattern (12). We used Pearson correlation as the clustering metric with complete linkage. The data for co-clustering with canonical chromatin states was downloaded from the supplemental information in Ref. 13. The figures were drawn using the tree map module of Processing visualization software (www.processing.org).

RESULTS AND DISCUSSION

We have devised a system that measures the performance of an antibody directed toward a specific histone post-translational modification (the "mark-specific" antibody) relative to the performance of a nonspecific antibody directed against the histone in general (the "pan-specific" antibody; Fig. 1*A*). In this manner we can: 1) measure the fold enrichment over background for the intended target of the mark-specific antibody, 2) measure the bias of the mark-specific antibody for or against any other potential post-translational modifications at the same site on the histone protein, and 3) estimate the background levels of discrete post-translational modifications at a site by using the observed abundance in the mass spectrometer.

This is accomplished by the use of SILAC technology to isotopically label histone proteins in cell culture (14). Histones are then isolated separately from each SILAC-labeled (light, medium, or heavy) cell population and chromatographically purified by reversed phase HPLC. One population, either the medium or heavy histones, is subjected to immunoprecipitation by the mark-specific antibody, whereas the other population (the light histones) is immunoprecipitated by the pan-specific antibody. After the separate immunoprecipitation experiments, the eluates are mixed together at a ratio of 1:1 by histone protein concentration. The resulting sample is separated by SDS-PAGE, and a molecular weight band corresponding to this histone of interest is excised. The intact proteins in the gel are then propionylated, digested into peptides, reproprionylated at the peptide level, and analyzed by LC-MS (Fig. 1*A*). The original sample source of the peptides observed in the mass spectrometer (either the mark-specific or pan-specific IP) is easily determined because of the shift in mass introduced by the atoms incorporated in the arginine residues of the histones. We chose to use purified histones for this study so that we could minimize confounding effects from higher order chromatin complexes.

As a concrete example, let us consider the evaluation of a mark-specific antibody directed toward H3K4me3 (Fig. 1*B*). In this case, we would incubate the mark-specific antibody with heavy H3 that contains $^{13}C_6$,$^{15}N_4$-arginine and the pan-specific antibody with light H3 that contains normal $^{12}C_6$,$^{14}N_4$-arginine. After the eluates of the IPs are mixed 1:1 by histone protein concentration and the H3 band is isolated from an SDS-PAGE experiment, the proteins are digested with trypsin. The propionylation steps ensure that trypsin will only cleave C-terminal to arginine. Thus, one expected product of digestion is a peptide spanning residues 3–8 of H3: T(K_{me3})QTAR. However, the peptide originating from the mark-specific IP will be ~10 Daltons heavier than its pan-specific counterpart (1 Arg-~10 additional Daltons in this example). Signals from these peptides are monitored in the course of the LC-MS experiment. Because the peptides derived from the heavy and light preparations are virtually chemically identical with respect to biophysical properties, they co-elute with reversed phase chromatography. The chromatographic area under the curve of traces corresponding to the MS signal of the two specific *m/z* values of the peptides can be used to compute the relative enrichment ratio of the targeted histone mark between the mark-specific and pan-specific antibodies.

We show in Fig. 1*B* how multiple possible modification states at H3K4 may be monitored in a single LC-MS experiment by generating XICs from their associated precursor ions. We additionally monitor an H3 peptide from residues 41–49 to serve as a proper normalization control. Two representations of the data are shown in the *lower right panel* of Fig. 1*B*. The *tree map* shows a box that is divided into segments with areas proportional to the amount of ion current from each of the H3K4 modification states from the pan-H3 enrichment (*blue*

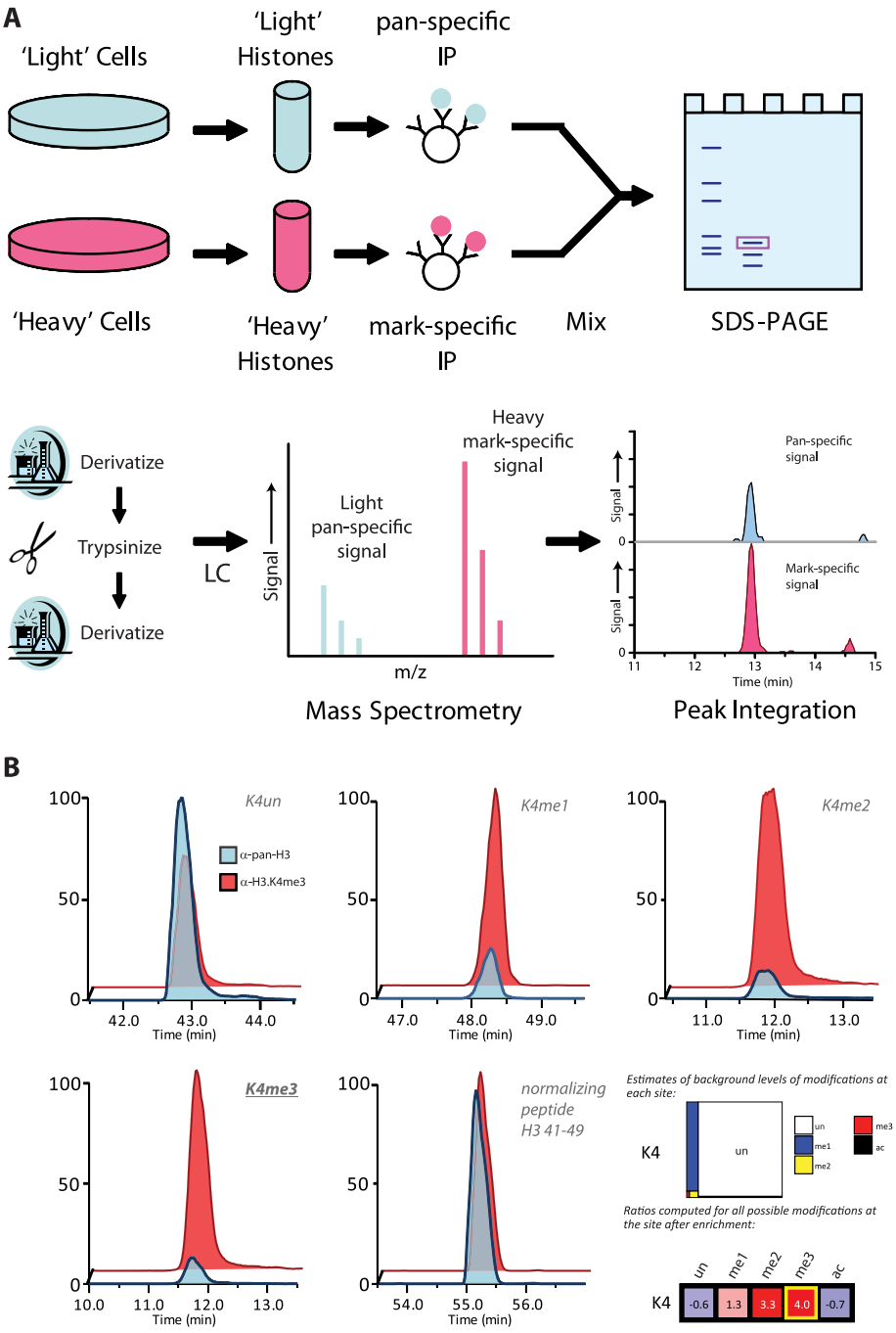

FIG. 1. **Antibody evaluation experimental paradigm and method of quantification.** *A*, histones were acid-extracted from HeLa S3 cells labeled in light or heavy SILAC growth medium, and H3 was further purified by reversed phase chromatography. Light H3 was incubated with a pan-specific H3 antibody, whereas heavy H3 was incubated with a mark-specific antibody under ChIP buffer conditions. The eluates were mixed and subjected to SDS-PAGE. The H3 band was derivatized with propionic anhydride, digested with trypsin, and then rederivatized (see "Experimental Procedures" for details). The peptides were analyzed by LC-MS on an Orbitrap MS instrument. *B*, this example demonstrates how H3K4 mark-targeting antibodies were evaluated. XICs corresponding to all possible modification states of the peptide bearing H3K4 were made for heavy (*red traces*) and light (*blue traces*) variants of the peptide. Additionally, we made XICs for an unmodified peptide spanning residues 41–49 of H3. The area under the curve was calculated for each XIC. The heavy/light ratio of the H3(41–49) peptide was used to normalize all other heavy/light ratios, thus ensuring that the same amount of H3 was being analyzed in each experiment. In the *lower right panel*, we derive two types of information. Each sub-box in the tree map (*upper portion*) is proportional to the amount of ion current from any given modification state over the total of all ion current from all modification states at H3K4, analogous to a pie chart. The heat map (*lower portion*) shows the $\log_2$ of the enrichment ratio (heavy/light) of the antibody being tested for each possible modification state of H3K4. *un*, unmodified; *me1*, monomethyl; *me2*, dimethyl; *me3*, trimethyl; *ac*, acetyl.

traces). This is analogous to a pie chart, and we take this as an estimate of the background level of the modification in bulk histones. The heat map has color-coded squares that each represent the ratio of the *red* peak area to the *blue* peak area for each modification state, normalized by the H3 41–49 peptide ratio. The $\log_2$ of this ratio is printed within each box, and the intended enrichment target of the antibody has a *yellow perimeter*. These ratios are the enrichments of each mark at a given site by the modification specific antibody relative to the pan-specific H3 antibody. See the figure legend for further details.

The input into our assay is full-length protein of a single histone type only, *i.e.*, only HPLC-purified H3 is used in testing an antibody directed against a mark on H3 (see "Experimental Procedures"). As such, we do not expect higher order histone complexes such as H3-H4 dimers or nucleosomes to be present during the assay. The use of a single histone type for each assay ensures the specificity of quantitation at the given site without confounding contributions from noncovalently associated partner histones. In addition, this strategy affords us the opportunity to assess whether other marks at distal sites co-enrich with a targeted mark. Observation of co-enrichment could be interpreted in two ways. Let us imagine that an antibody directed against H3K4me3 is found to enrich both H3K4me3 and H3K36me3. In one interpretation, the antibody has affinity for a mark at a site other than the intended one (*i.e.*, the antibody against K4me3 also recognizes K36me3). In an alternative interpretation, the antibody enriches only H3K4me3, but histone molecules with this mark are more likely to have H3K36me3 than the background.

We performed all work in suspension-grown HeLa cells. We could easily observe many of the known modifications to H3 using this system. Outside of this work, we have profiled histone modifications in several other cell types of varying lineages. Although some lineage-specific differences in the background levels of different histone modifications do exist, the HeLa cells used herein appear to be generally representative of most other cells types that we have tested.

H3K4-targeting Antibodies—We examined several antibodies directed at post-translational modifications on H3K4 (Fig. 2*A*). H3K4me3 has canonically been associated with transcriptional start sites of active genes, and H3K4me2 is associated with promoter regions and enhancers of highly expressed genes (13, 15). Both are considered "active" chromatin marks. We chose ChIP grade antibodies against these targets for profiling.

CST9726 is directed against H3K4me2. This antibody enriched that target more than 16-fold with only minor enrichments from H3K4me1 and H3K4me3. It also biased against unmodified H3K4 and H3K4ac (Fig. 2*A*). We estimate that H3K4me2 composes less than 1% of all H3K4 in HeLa cells. The estimate of this quantity is derived by examining the proportion of ion current belonging to the H3K4me2-containing peptide (summing all charge states observed) in the sum

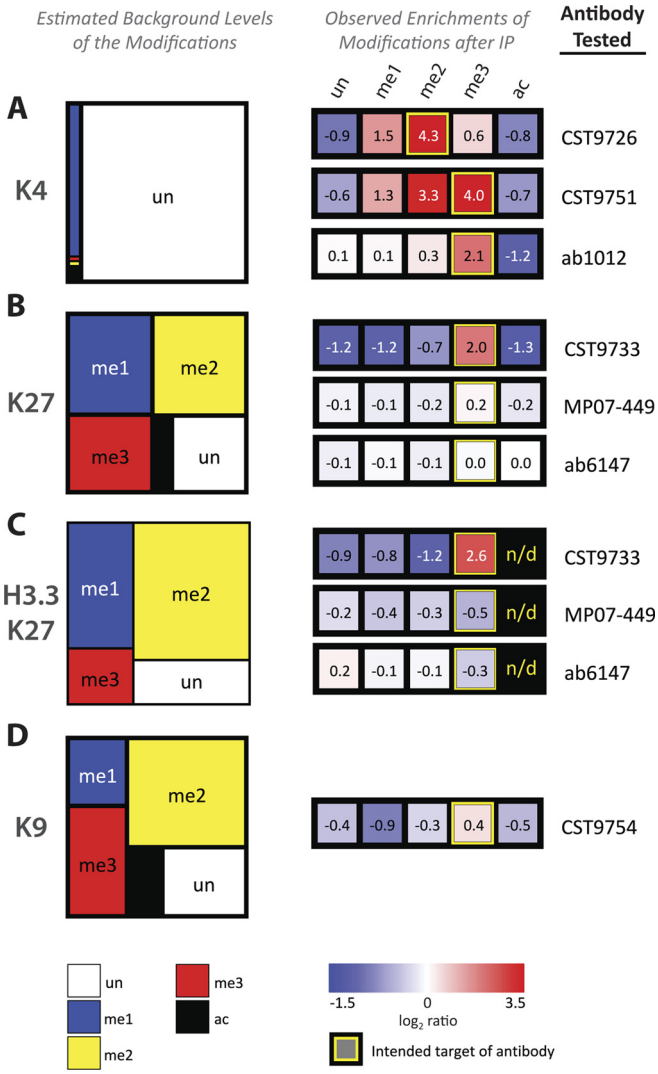

FIG. 2. **Evaluation of antibodies directed against H3K4, H3K27, and H3K9 modifications.** Tree maps and heat maps for each antibody evaluated are annotated as in the *lower right panel* of Fig. 1*B*. Note that the intended target of enrichment is bounded by a *yellow box* in each heat map. Where necessary, we collapsed several modification combinations into a single value to represent enrichment of the desired mark (*i.e.*, H3K9me3K14un + H3K9me3K14ac). This was done by computing an intensity-weighted average of all modification combinations that bear the mark stated in the column. The intensity of the pan-specific background was used to weight each combination. *A*, evaluation of H3K4 targeting antibodies. *B*, evaluation of H3K27me3 targeting antibodies on their ability to enrich marks on the H3.1 and H3.2 variants of H3 (Ala-31 variant). *C*, evaluation of H3K27me3 targeting antibodies on their ability to enrich marks on the H3.3 variant of H3 (Ser-31 variant). *D*, evaluation of an H3K9me3 targeting antibody.

of ion current derived from all modification states (again, across all charge states observed), namely: H3K4un, H3K4me1, H3K4me2, H3K4me3, and H3K4ac1, in the pan-specific antibody condition. These percentages are further normalized by empirical measurements of the ionization efficiencies of synthetic versions of these peptides (see "Exper-

imental Procedures" and supplemental Table 2). CST9726 would be a nearly ideal ChIP antibody, providing a high signal-to-noise ratio against a background composed of mostly other forms of H3.

We also tested two antibodies directed against H3K4me3 (Fig. 2*A*). CST9751 also provided a high level of enrichment for its intended target (~16-fold), but it also seemed to enrich another mark at the same site: H3K4me2, by nearly 10-fold. Because these two marks cannot be present on the same site at the same time and histones are not assembled into nucleosomes under the conditions of our assay (which might allow for this dual enrichment), we must conclude that CST9751 has significant off target enrichment of H3K4me2. The manufacturer does state that there may be some cross-reactivity with H3K4me2 for this antibody. However, we can use our method to objectively quantify this cross-reactivity and conclude that it is quite significant. The second antibody that we tested against H3K4me3 (ab1012) was less efficient at enriching H3K4me3 over background but more selective. It enriched H3K4me3 by over 4-fold but did not seem to enrich any other state, including H3K4me2, and it was even biased against H3K4ac1. We would consider this a suitable antibody for ChIP applications where true selectivity toward H3K4me3 was desired.

H3K27-targeting Antibodies—We tested several antibodies directed against H3K27me3 in the same manner (Fig. 2*B*). H3K27me3 is associated with repressed regions of the genome and is deposited by the Polycomb Repressive Complex 2 enzyme EZH2 (16). This enzyme has recently been shown to be of interest in some cancers (17, 18). Study of H3K27 carries an additional challenge in that it is located on the same peptide as another PTM-bearing residue, H3K36. Therefore, to fully examine H3K27, we also needed to consider all of the concomitant states of H3K36. The enrichment ratios that we report are computed by weighting each state (*i.e.*, H3K27me3/H3K36un or H3K27me3/H3K36me2) by its ion current relative to the total of all states bearing the same H3K27 modification. We note that our chromatography was able to separate isobaric peptides (*i.e.*, H3(27-40) in the H3K27me2K36*un* H3K27*un*K36me2 states), and therefore precursor level quantification should be valid for discerning such species.

We should first consider the background frequencies of the various modifications at H3K27. According to our estimates (Fig. 2*B*, *tree map*), H3K27 exists most commonly in the dimethylated state (~30%). We have found this to be true in murine embryonic stem cells and murine embryonic fibroblasts as well (data not shown). H3K27me3 composes almost ~21% of the population, whereas H3K27un is only ~18% of the population. This is in stark contrast to the situation at H3K4, where the canonical active mark composes less than 1% of the population. By these estimates, we can conclude that the canonical "repressive mark" is distributed much more widely in the genome than the active mark. Moreover, a significant population is poised to be trimethylated by the addition of a single additional methyl group. The functional significance of H3K27me2 has not been systematically assessed.

Three antibodies directed against H3K27me3 were tested: CST9733, MP07-449, and ab6147 (Fig. 2*B*). In our hands, CST9733 was vastly superior to the other antibodies. CST9733 enriches H3K27me3 ~4-fold over background and also actively biases against H3K27 in any other state. These are ideal properties for a ChIP grade antibody. In contrast, MP07-449 (which is widely used in ChIP studies) only marginally enriches of H3K27me3 (~1.1-fold) but also biases against other forms of H3K27 (importantly, H3K27me2, which we have shown is likely the most abundant state). Therefore we would expect this antibody to have poor signal-to-noise properties in a ChIP assay. Ab6147 does not seem to enrich for the intended target at all, although we cannot exclude its usefulness in Western blot experiments from these data.

Use of mass spectrometry can, in this instance, further dissect antibody selectivity properties. The replication-dependent variants of H3, H3.1, and H3.2 contain an Ala residue at position 31, whereas the replication-independent variant H3.3 contains a Ser residue at this position. Because this amino acid residue is contained within the peptide used to monitor H3K27, we can discern whether there are any biases of the antibodies for or against these variants. Indeed, Fig. 2*C* shows that CST9733 is insensitive to the amino acid at position 31. However, both of the other antibodies actively bias against the form containing Ser-31 regardless of the state. In this sense, MP07-449 could be said to be H3.1/2-selective to a certain degree, whereas CST9733 would give a researcher an unbiased assessment of the state of H3K27me3 regardless of histone variant.

H3K9-targeting Antibodies—Thus far we have evaluated one H3K9me3-targeting antibody, CST9754 (Fig. 2*D*). H3K9me3 is another canonical mark of genomic repression. Again, it is useful to note that the most prevalent state of H3K9 is in the dimethylated form (~41%), and we estimate that H3K9me3 composes ~21% of the population. CST9754 somewhat enriches H3K9me3 (~1.3-fold) and somewhat biases against other forms, notably 1.2-fold against H3K9me2. We did not measure the effect of a proximal phosphorylation at H3S10 on the performance of the antibodies because we did not generally observe this modification in any state; therefore any biases against the presence of this mark would be negligible in a ChIP experiment. We may evaluate more reagents against this mark in the future.

Background Frequencies of Other Histone Modifications on H3—We also estimated the total amounts of other well known histone modifications on H3 in HeLa using the same techniques as employed above (Fig. 3). We found relatively high background levels of H3K14ac and H3K23ac. H3K18ac was noticeably less abundant than these other commonly acety-

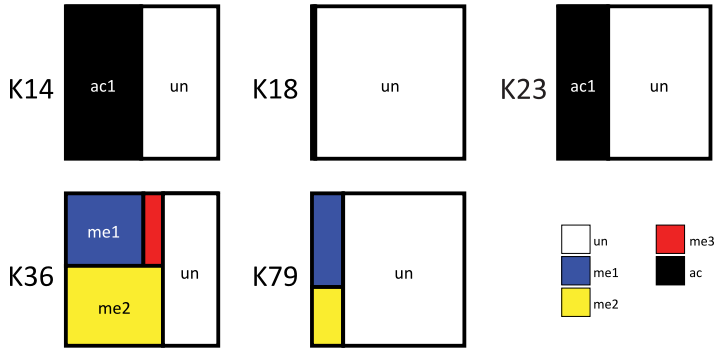

FIG. 3. **Estimation of background frequencies of histone modifications at other sites.** Background frequencies were estimated using the techniques detailed under "Experimental Procedures" and are depicted in the tree map format explained in the legend for Fig. 1*B*.

lated residues. H3K36 showed a high degree of mono- and dimethylation (~24 and ~33%, respectively), although the unmodified state was still the most common (~36%). H3K36 acetylation was not observed in these cells. Numerical representations of all background frequencies can be found in supplemental Table 3.

Co-occurrence of Modifications on Histone H3—We did not limit ourselves to measuring enrichment of histone modifications solely at the sites targeted by the antibodies that we characterized. We also measured enrichment ratios at a variety of observable sites on H3. We then subjected these data to hierarchical clustering of both the rows and columns as shown in Fig. 4*A*. It was immediately apparent that data from multiple antibodies targeting active marks (*e.g.* H3K4me2, H3K4me3) clustered together and that data from multiple antibodies targeting repressive marks (*e.g.* H3K9me3, H3K27me3) also clustered together. The data from the pan-specific H3 antibody (when compared with itself) fell on the boundary between these two groups.

We also noted that formation of these clusters was being driven by marks in addition to those that were targeted by the antibodies. For example, our experiments showed co-enrichment of the simultaneous acetylation of H3K18 and H3K23 in conjunction with antibody enrichment of H3K4me2. The H3K18 and H3K23 sites fall on the same peptide in our experiment and can be monitored together. Monitoring of this doubly acetylated species is independent of monitoring of either the H3K18 or H3K23 singly acetylated species. Thus, when H3K4me2 is present, H3K18ac and H3K23ac are likely to also be present on the same histone molecule. H3K79me1 and H3K79me2 were likewise co-enriched with H3K4me2. Depletion of H3K27me3 in the presence of H3K4me2 was also observed. Corroborating these observations, Zhang *et al.* (19) noticed that overall levels of both K4 and K79 methylation increased as the acetylation state of H3 increased.

Analogously, H3K27me3 and H3K9me3 antibodies co-enriched the second repressive mark along with the targeted one, *i.e.*, CST9733 which targets H3K27me3 co-enriches H3K9me3. Both also enrich for the unmodified states at H3K4 and H3K36, both considered sites of activating modifications. However, marks at other sites did not mutually reinforce these repressive clusters. Enrichment of H3K27me3 comes with a concomitant depletion of H3K79 methylation, whereas H3K9me3 enrichment co-enriches H3K79 methylation states.

Co-occurrence of marks at specific loci within the genome has sometimes been referred to as a histone code that may control the transcriptional output of a genomic region (1). The development of microarrays and massively parallel sequencing technologies as readouts for ChIP experiments has enabled generation of a wealth of data that map specific histone marks onto particular loci of the genome in various cell types (20, 21). These studies generally support the notion that multiple marks may be simultaneously present at a given locus, and these combinations can be correlated with the transcriptional state of the locus.

As stated earlier, the inputs into our immunoaffinity enrichment experiments are HPLC-purified histone H3 populations, in isolation from other core components of the nucleosome octamer. We do not expect higher order protein-protein interactions to occur, and histone molecules should be monomeric in the assay. Therefore, we might draw one of two conclusions from these observations of co-enrichment of marks. The trivial explanation is that the antibodies in question actually have significant off target activities at distal sites from the intended target. The alternative explanation, and the one that we favor here, is that these histone codes are encoded on single histone protein molecules.

One drawback of our methodology is that we cannot easily distinguish between these two possibilities. If we had access to populations of recombinant histones, each bearing only a single defined modification, we might be able to do so. Another possibility would be to use synthetic peptides in our experiment with defined modifications to attempt to "compete" off histone proteins in an epitope-specific manner. This route would be costly and, without knowledge of the original epitopes to which the antibodies were raised, could involve a good deal of trial and error.

However, there is ample existing evidence that supports our latter hypothesis. For example, several studies conducted on large fragments of H3 (residues 1–50) have demonstrated the existence of all of these potential combinations in cells (22–24). Although no enrichments were performed in these prior studies to establish the direct correlation between one

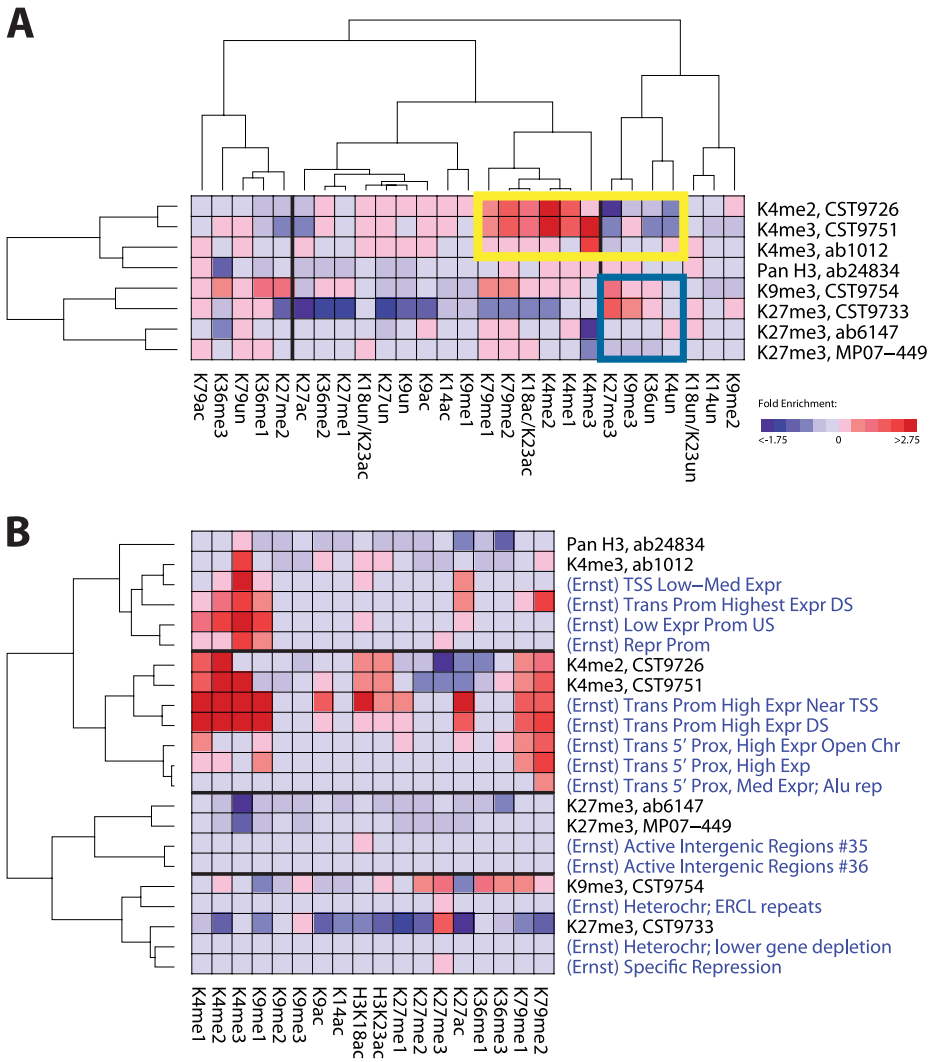

FIG. 4. **Hierarchical clustering of antibody enrichment histone modification data.** *A*, enrichment ratios for all of the sites measured on H3 were clustered (Pearson correlation, complete linkage) by modification and antibody enrichment. Each *row* corresponds to a particular antibody enrichment strategy and is labeled with the intended target of the antibody and its catalogue number. Each *column* corresponds to a different histone modification, where *un* means unmodified. Each *square* is color-coded by an intensity-weighted average of all modification combinations that bear the mark stated in the column. Intensity of the pan-specific background was used to weight each combination. The *yellow box* depicts a core set of modification enrichments that drive clustering of antibodies together that target canonical activating marks. The *blue box* depicts a core set of modification enrichments that drive clustering of antibodies together that target canonical repressive marks. *B*, antibody enrichment data were clustered together with canonical chromatin states as determined computationally by Ernst and Kellis (13). Only rows were clustered in this case, and the data set was limited to histone modifications that were common to the two studies. Some of Ernst's states are omitted for clarity. Because the data sets are on two different scales, the color scale is arbitrary. See text for details.

mark and another, we can say that they at least exist in nature. In principle, the same enrichments performed in these experiments could be analyzed by the "middle-down" approach targeting the H3 1–50 fragment. We have attempted such studies in our laboratory and found that the amount of antibody required to generate enough input material would be not be feasible.

We also note that the co-enrichment patterns recapitulate a great deal of established functional biology. Antibodies targeted against marks of active chromatin (*i.e.*, H3K4me2 or H3K4me3) seem to co-enrich other marks of active chromatin (H3K18ac and H3K79me2) and vice versa. A further case against antibody cross-reactivity is that an antibody would need to be promiscuous in such a way in that it would enrich modification states that correspond with known biology derived from alternative methods.

Ernst and Kellis (13) recently reported a set of canonical chromatin states derived from a meta-analysis of multiple ChIP data sets from T-cells. Many of these states could be mapped to recognizable defined elements in the genome. These canonical states were meant to reflect combinations of histone marks present at generic loci in the genome with

known biological interpretations. For example, Ernst and Kellis determined the combinations of histone modifications likely to be present in the 5′-region of a gene locus that is highly expressed. These combinations would be one class of canonical state. We extended our analysis by clustering our orthogonal proteomics data with the reported ChIP-based canonical states when there were modifications in common between the two studies (Fig. 4*B*).

Somewhat to our surprise, data from the IPs (*black row labels* in Fig. 4*B*) intercalated well with Ernst's canonical states (*blue row labels* in Fig. 4*B*) in the resulting dendrogram. We noted that multiple marks contributed to the formation of clusters, resulting in strong overlap between the enriched H3 and the canonical states. This may suggest that the canonical states themselves are represented on single histone molecules, rather than as an ensemble of singly modified histones present at a general locus. Given the separation between antibodies ostensibly targeting the same mark, our data also suggest that certain antibodies are better considered as proxies for canonical chromatin states rather than specific targets of given modifications.

Summation—We have developed a method that allows the quantitative profiling and characterization of antibodies that are used in ChIP assays to study post-translational modifications of histone proteins. Knowledge of the properties of these reagents could be of great benefit when designing and interpreting ChIP experiments. For example, an *ad hoc* examination of H3K27me3 ChIP "tracks" obtained using the MP07-449 antibody shows diffuse H3K27me3 regions with highly variability in H3K27me3 level. Our studies show that MP07-449 is a relatively poor performer, which might help to explain some of these observations. We also estimate that the global H3K27me3 background is quite high (~20%), which would set an upper limit of signal-to-noise of about 5:1. Although local backgrounds of H3K27me3 may differ at specific loci, knowing the global background could allow better models of defining H3K27me3-containing regions. We hope that the data presented here will be useful to the epigenetics community as more ChIP experimental data are obtained.

The use of mass spectrometry to study histone modifications is orthogonal to ChIP experiments. MS experiments are capable of monitoring multiple histone marks in a single experiment albeit with no genomic resolution, whereas ChIP-Seq experiments can only probe a single mark at a time with far better genomic resolution. Using our system, we found that certain combinations of histone marks were co-enriched on the same histone H3 protein molecule. Remarkably, these combinations largely reinforce the current functional interpretations of the posited histone code. Indeed, the unexpected overlap of our co-enrichment patterns with the canonical chromatin states derived from ChIP experiments suggest that the code may be "written" on single histone proteins. We hope that these observations can be independently verified and can provide novel ideas in epigenetics research. For example, observation of multiple functionally reinforcing histone marks on the same protein raises a molecular "chicken and the egg " question as to which of the marks appears first and how the subsequent marks follow. Are there stepwise dependences that dictate the accumulation of the fully explicated code on a single molecule? How does the code apply at the level of a single nucleosome? We could use our technique to enrich preparations of mononucleosomes rather than purified histone proteins to study this latter question. We believe that we have demonstrated how MS techniques can synergize with ChIP techniques to generate a comprehensive assessment of the functional chromatin landscape and further our understanding of epigenetic mechanisms.

Acknowledgments—We thank Dr. Bang Wong of the Broad Institute for useful discussions on visualization of the data herein. We also thank Dr. Jason Ernst and Dr. Manolis Kellis of the Massachusetts Institute of Technology for helpful discussions regarding the data from their paper.

* This work was supported in part by National Institutes of Health Grant R21 DA025720 (to J. J.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

S This article contains supplemental material.

‡ To whom correspondence should be addressed: The Broad Institute, Proteomics Platform, 7 Cambridge Center, Rm. 5025, Cambridge, MA 02142. Tel.: 617-714-7638; Fax: 617-714-8957; E-mail: jjaffe@broadinstitute.org.

REFERENCES

1. Jenuwein, T., and Allis, C. D. (2001) Translating the histone code. *Science* **293,** 1074–1080
2. Solomon, M. J., Larsen, P. L., and Varshavsky, A. (1988) Mapping protein-DNA interactions *in vivo* with formaldehyde: Evidence that histone H4 is retained on a highly transcribed gene. *Cell* **53,** 937–947
3. Solomon, M. J., and Varshavsky, A. (1985) Formaldehyde-mediated DNA-protein crosslinking: A probe for *in vivo* chromatin structures. *Proc. Natl. Acad. Sci. U.S.A.* **82,** 6470–6474
4. Martin, C., and Zhang, Y. (2005) The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* **6,** 838–849
5. Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L., and Lander, E. S. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125,** 315–326
6. Eberl, H. C., Mann, M., and Vermeulen, M. (2011) Quantitative proteomics for epigenetics. *Chembiochem* **12,** 224–234
7. Thomas, C. E., Kelleher, N. L., and Mizzen, C. A. (2006) Mass spectrometric characterization of human histone H3: A bird's eye view. *J. Proteome Res.* **5,** 240–247
8. Kinter, M., and Sherman, N. E. (2000) *Protein Sequencing and Identification Using Tandem Mass Spectrometry*, Wiley-Interscience, New York
9. Garcia, B. A., Mollah, S., Ueberheide, B. M., Busby, S. A., Muratore, T. L., Shabanowitz, J., and Hunt, D. F. (2007) Chemical derivatization of histones for facilitated analysis by mass spectrometry. *Nat. Protoc.* **2,** 933–938
10. Jaffe, J. D., Keshishian, H., Chang, B., Addona, T. A., Gillette, M. A., and Carr, S. A. (2008) Accurate inclusion mass screening: A bridge from unbiased discovery to targeted assay development for biomarker verification. *Mol. Cell. Proteomics* **7,** 1952–1962
11. MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., and MacCoss, M. J. (2010) Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26,** 966–968

12. Reich, M., Liefeld, T., Gould, J., Lerner, J., Tamayo, P., and Mesirov, J. P. (2006) GenePattern 2.0. *Nat. Genet.* **38,** 500–501
13. Ernst, J., and Kellis, M. (2010) Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol.* **28,** 817–825
14. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1,** 376–386
15. Pekowska, A., Benoukraf, T., Zacarias-Cabeza, J., Belhocine, M., Koch, F., Holota, H., Imbert, J., Andrau, J. C., Ferrier, P., and Spicuglia, S. (2011) H3K4 tri-methylation provides an epigenetic signature of active enhancers. *EMBO J.* **16,** 4198–4210
16. Bracken, A. P., and Helin, K. (2009) Polycomb group proteins: Navigators of lineage pathways led astray in cancer. *Nat. Rev. Cancer* **9,** 773–784
17. Ernst, T., Chase, A. J., Score, J., Hidalgo-Curtis, C. E., Bryant, C., Jones, A. V., Waghorn, K., Zoi, K., Ross, F. M., Reiter, A., Hochhaus, A., Drexler, H. G., Duncombe, A., Cervantes, F., Oscier, D., Boultwood, J., Grand, F. H., and Cross, N. C. (2010) Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat. Genet.* **42,** 722–726
18. Morin, R. D., Johnson, N. A., Severson, T. M., Mungall, A. J., An, J., Goya, R., Paul, J. E., Boyle, M., Woolcock, B. W., Kuchenbauer, F., Yap, D., Humphries, R. K., Griffith, O. L., Shah, S., Zhu, H., Kimbara, M., Shashkin, P., Charlot, J. F., Tcherpakov, M., Corbett, R., Tam, A., Varhol, R., Smailus, D., Moksa, M., Zhao, Y., Delaney, A., Qian, H., Birol, I., Schein, J., Moore, R., Holt, R., Horsman, D. E., Connors, J. M., Jones, S., Aparicio, S., Hirst, M., Gascoyne, R. D., and Marra, M. A. (2010) Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat. Genet.* **42,** 181–185
19. Zhang, K., Siino, J. S., Jones, P. R., Yau, P. M., and Bradbury, E. M. (2004) A mass spectrometric "Western blot" to evaluate the correlations between histone methylation and histone acetylation. *Proteomics* **4,** 3765–3775
20. Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell* **129,** 823–837
21. Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigó, R., Gingeras, T. R., Margulies, E. H., Weng, Z., Snyder, M., Dermitzakis, E. T., Thurman, R. E., Kuehn, M. S., Taylor, C. M., Neph, S., Koch, C. M., Asthana, S., Malhotra, A., Adzhubei, I., Greenbaum, J. A., Andrews, R. M., Flicek, P., Boyle, P. J., Cao, H., Carter, N. P., Clelland, G. K., Davis, S., Day, N., Dhami, P., Dillon, S. C., Dorschner, M. O., Fiegler, H., Giresi, P. G., Goldy, J., Hawrylycz, M., Haydock, A., Humbert, R., James, K. D., Johnson, B. E., Johnson, E. M., Frum, T. T., Rosenzweig, E. R., Karnani, N., Lee, K., Lefebvre, G. C., Navas, P. A., Neri, F., Parker, S. C., Sabo, P. J., Sandstrom, R., Shafer, A., Vetrie, D., Weaver, M., Wilcox, S., Yu, M., Collins, F. S., Dekker, J., Lieb, J. D., Tullius, T. D., Crawford, G. E., Sunyaev, S., Noble, W. S., Dunham, I., Denoeud, F., Reymond, A., Kapranov, P., Rozowsky, J., Zheng, D., Castelo, R., Frankish, A., Harrow, J., Ghosh, S., Sandelin, A., Hofacker, I. L., Baertsch, R., Keefe, D., Dike, S., Cheng, J., Hirsch, H. A., Sekinger, E. A., Lagarde, J., Abril, J. F., Shahab, A., Flamm, C., Fried, C., Hackermüller, J., Hertel, J., Lindemeyer, M., Missal, K., Tanzer, A., Washietl, S., Korbel, J., Emanuelsson, O., Pedersen, J. S., Holroyd, N., Taylor, R., Swarbreck, D., Matthews, N., Dickson, M. C., Thomas, D. J., Weirauch, M. T., Gilbert, J., Drenkow, J., Bell, I., Zhao, X., Srinivasan, K. G., Sung, W. K., Ooi, H. S., Chiu, K. P., Foissac, S., Alioto, T., Brent, M., Pachter, L., Tress, M. L., Valencia, A., Choo, S. W., Choo, C. Y., Ucla, C., Manzano, C., Wyss, C., Cheung, E., Clark, T. G., Brown, J. B., Ganesh, M., Patel, S., Tammana, H., Chrast, J., Henrichsen, C. N., Kai, C., Kawai, J., Nagalakshmi, U., Wu, J., Lian, Z., Lian, J., Newburger, P., Zhang, X., Bickel, P., Mattick, J. S., Carninci, P., Hayashizaki, Y., Weissman, S., Hubbard, T., Myers, R. M., Rogers, J., Stadler, P. F., Lowe, T. M., Wei, C. L., Ruan, Y., Struhl, K., Gerstein, M., Antonarakis, S. E., Fu, Y., Green, E. D., Karaöz, U., Siepel, A., Taylor, J., Liefer, L. A., Wetterstrand, K. A., Good, P. J., Feingold, E. A., Guyer, M. S., Cooper, G. M., Asimenos, G., Dewey, C. N., Hou, M., Nikolaev, S., Montoya-Burgos, J. I., Löytynoja, A., Whelan, S., Pardi, F., Massingham, T., Huang, H., Zhang, N. R., Holmes, I., Mullikin, J. C., Ureta-Vidal, A., Paten, B., Seringhaus, M., Church, D., Rosenbloom, K., Kent, W. J., Stone, E. A., Batzoglou, S., Goldman, N., Hardison, R. C., Haussler, D., Miller, W., Sidow, A., Trinklein, N. D., Zhang, Z. D., Barrera, L., Stuart, R., King, D. C., Ameur, A., Enroth, S., Bieda, M. C., Kim, J., Bhinge, A. A., Jiang, N., Liu, J., Yao, F., Vega, V. B., Lee, C. W., Ng, P., Yang, A., Moqtaderi, Z., Zhu, Z., Xu, X., Squazzo, S., Oberley, M. J., Inman, D., Singer, M. A., Richmond, T. A., Munn, K. J., Rada-Iglesias, A., Wallerman, O., Komorowski, J., Fowler, J. C., Couttet, P., Bruce, A. W., Dovey, O. M., Ellis, P. D., Langford, C. F., Nix, D. A., Euskirchen, G., Hartman, S., Urban, A. E., Kraus, P., Van Calcar, S., Heintzman, N., Kim, T. H., Wang, K., Qu, C., Hon, G., Luna, R., Glass, C. K., Rosenfeld, M. G., Aldred, S. F., Cooper, S. J., Halees, A., Lin, J. M., Shulha, H. P., Zhang, X., Xu, M., Haidar, J. N., Yu, Y., Iyer, V. R., Green, R. D., Wadelius, C., Farnham, P. J., Ren, B., Harte, R. A., Hinrichs, A. S., Trumbower, H., Clawson, H., Hillman-Jackson, J., Zweig, A. S., Smith, K., Thakkapallayil, A., Barber, G., Kuhn, R. M., Karolchik, D., Armengol, L., Bird, C. P., de Bakker, P. I., Kern, A. D., Lopez-Bigas, N., Martin, J. D., Stranger, B. E., Woodroffe, A., Davydov, E., Dimas, A., Eyras, E., Hallgrímsdóttir, I. B., Huppert, J., Zody, M. C., Abecasis, G. R., Estivill, X., Bouffard, G. G., Guan, X., Hansen, N. F., Idol, J. R., Maduro, V. V., Maskeri, B., McDowell, J. C., Park, M., Thomas, P. J., Young, A. C., Blakesley, R. W., Muzny, D. M., Sodergren, E., Wheeler, D. A., Worley, K. C., Jiang, H., Weinstock, G. M., Gibbs, R. A., Graves, T., Fulton, R., Mardis, E. R., Wilson, R. K., Clamp, M., Cuff, J., Gnerre, S., Jaffe, D. B., Chang, J. L., Lindblad-Toh, K., Lander, E. S., Koriabine, M., Nefedov, M., Osoegawa, K., Yoshinaga, Y., Zhu, B., and de Jong, P. J. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447,** 799–816
22. Garcia, B. A., Pesavento, J. J., Mizzen, C. A., and Kelleher, N. L. (2007) Pervasive combinatorial modification of histone H3 in human cells. *Nat. Methods* **4,** 487–489
23. Taverna, S. D., Ueberheide, B. M., Liu, Y., Tackett, A. J., Diaz, R. L., Shabanowitz, J., Chait, B. T., Hunt, D. F., and Allis, C. D. (2007) Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 2086–2091
24. Young, N. L., DiMaggio, P. A., Plazas-Mayorca, M. D., Baliban, R. C., Floudas, C. A., and Garcia, B. A. (2009) High throughput characterization of combinatorial histone codes. *Mol. Cell. Proteomics* **8,** 2266–2284